Emergent Structure in Dense Packings as Informational Tunneling (q→τ): Local Flips, Avalanches, Absorbing States—& Parallels to Learning by Compression

Douglas C. Youvan

doug@youvan.com

www.youvan.ai

October 19, 2025

Dense 2D packings self-organize without hand-coded rules. Near jamming, local neighbor swaps (T1), shear-like plastic events, and avalanches rewire the contact graph in discrete bursts that appear instantaneous on the protocol clock. We cast these bursts as informational tunneling along a single alignment path q→τ: the realized microstate q (contacts, defects, spectra) descends a declared misfit D(q‖τ) toward a constraint-defined manifold τ(isostaticity, hexatic or hyperuniform statistics, absorbing/no-collision sets). An update is admissible precisely when it monotonically reduces Don the chosen observables. Under quasi-static densification or weak cyclic drive the system exhibits plateaus and step drops in $D(t)$, mobile defect "packets," and approach to stationary sets with $\nabla_q D = 0$. Crucially, heterogeneity enters only through a distribution of sizes/shapes; the effective local "rules" are induced by mechanics acting on that distribution. We also outline an operational parallel to learning by compression in large language models: staircase loss descent, discrete representation phase changes, and a bias toward compressed, modular structure mirror the packing dynamics (without asserting mechanistic identity). Our contribution is a unified q→τ account, diagnostics (flip-field rasters, motif-KL, $S(k)$-KL, defect kinematics, transfer entropy), and minimal protocols that demonstrate computation from a distribution—in packings and, by analogy, in learned representations.

Keywords: jamming, T1 transitions, shear transformation zones, avalanches, random organization, absorbing states, hyperuniformity, hexatic order, contact graph, defect kinematics, informational tunneling, misfit D, q→τ, admissible descent, computation from a distribution, learning by compression, LLMs, representation phase changes.

A collaboration with GPT-5. CC4.0.

Outline

1. Premise: Computation from a Distribution
   1.1 Rules vs distributions • 1.2 Packings as mechanics-only processors
2. Framework
   2.1 Definitions: q, τ, observable set, $D(q\|\tau)$
   2.2 Process principle: one path, multiple clocks (time, iteration)
   2.3 Admissible descent: $\Delta D \leq 0$on declared observables
3. Local Events as Tunneling
   3.1 T1/STZ as single-site barrier crossings
   3.2 Avalanches as multi-site tunneling; staircase $D(t)$
   3.3 Absorbing τunder cyclic drive (random organization)
4. Targets (τ) & Metrics (D)
   4.1 Isostatic closures and contact number
   4.2 Hexatic/defect statistics (ψ_6, 5–7 pairs)
   4.3 Spectral targets $S(k)$ (hyperuniformity)
   4.4 Motif libraries; KL on motif frequencies
5. Induced Local Logic from a Distribution
   5.1 Uniform disks; doped shapes/size mixtures
   5.2 Post-hoc “rule tables” via neighborhood motifs
   5.3 Phase diagram over moments of the distribution
6. Protocols & Diagnostics
   6.1 Quasi-static compression / weak periodic drive with PBC
   6.2 Flip-field rasters, defect tracking, transfer entropy
   6.3 Reproducible staircases; hysteresis tests
7. Parallels to Learning by Compression (LLMs) — *comparative, not mechanistic*
   7.1 Mapping: $q, \tau, D \leftrightarrow$ parameters, data/prior, loss/KL
   7.2 Plateau-and-jump descent; representation phase changes
   7.3 Emergent modularity ↔ hyperuniform/defect-templated order
   7.4 Caveats and limits of the analogy
8. Results Archetypes
   8.1 Glider-like defect packets and collisions
   8.2 Convergence to absorbing sets; random organization
   8.3 Emergent hyperuniformity without explicit rules

9. Discussion & Outlook
   9.1 Consistency with prior q→τ work
   9.2 Open problems (proofs of monotonicity, optimal observables, 3D/finite-T)
   9.3 Designing computation by choosing distributions, not rules

References

1. Premise: Computation from a Distribution

1.1 Rules vs distributions

Classical cellular automata hard-code local update rules: a lookup table maps a neighborhood's finite state to a new state. By contrast, computation from a distribution delegates "the rule" to physics acting on microscopic heterogeneity sampled from a distribution. In our case the only ingredients are: (i) non-overlap (hard interactions) and deterministic relaxation; (ii) a protocol clock (quasi-static compression or weak cyclic drive); and (iii) a distribution over microscopic attributes (e.g., sizes or shapes). There is no hand-written logic. Yet, as the system evolves, it exhibits effective local logic—which neighbor contacts flip, which defects move, when cascades trigger—purely as a consequence of how mechanics couples to the statistics drawn from that distribution.

We formalize this with q→τ. Let q denote the realized microstate (contacts, defects, spectra). Let τ denote a constraint-defined manifold (e.g., isostaticity, hexatic or hyperuniform statistics, absorbing/no-collision sets) expressed via an observable set. A declared misfit $D(q \parallel \tau)$ provides the direction of improvement. An update is admissible if and only if it makes D decrease on the declared observables. Under this lens, a local neighbor swap (T1) is not "applied by a rule"; it is selected by the confluence of geometry, forces, and the sampled heterogeneity—an entropic tunneling step along the same alignment path. The distribution fixes the menu of motifs (e.g., size triplets, shape interlocks) and their frequencies; mechanics turns those motifs into propensities for specific local updates. Changing only the distribution changes the induced "rule table," while the physics and admissibility criterion remain fixed.

This yields a rigorous separation: rules are explicit, externally specified maps; distributions seed induced local logic through admissible descent of D. The latter supports programmability without writing rules: choose a distribution; the system computes.

1.2 Packings as mechanics-only processors

A packing near jamming is a mechanics-only processor: inputs are draws from a distribution (sizes, shapes, orientations); the processing unit is geometric compatibility plus relaxation; outputs are spatiotemporal patterns of flips, defect motion, and spectral organization. The computation performs informational tunneling: long plateaus where q is trapped in a metastable basin, punctuated by discrete, localized barrier crossings that drop D. On the protocol clock these appear instantaneous, yielding a staircase $D(t)$. Because updates are local and propagate through elastic/constraint couplings, the system naturally supports signal carriers—defect packets—that travel, interact, and annihilate, enabling CA-like composition of operations without external control.

The same device can target different τ by changing only the observable set in D. For example:

- Isostatic τ: monitor contact number and residual overlaps; admissible updates close constraints.
- Hexatic τ: monitor | ψ_6 |statistics and 5–7 defect densities; admissible updates reduce free disclinations and bind dislocations.
- Hyperuniform τ: monitor small-k structure factor $S(k)$; admissible updates suppress long-wavelength number fluctuations.
- Absorbing τ (random organization): monitor event/collision rate; admissible updates move the system toward no-future-event states under a fixed drive.

Critically, programmability resides in the distribution, not in bespoke dynamics. A narrow monodisperse distribution induces crystalline order with rare, easily healed defects; a finely tuned bimodal ratio promotes frustration that spawns mobile defect lines (our computational "wires"); doping by shape (e.g., sparse equilateral triangles among disks) templates junctions and turns on specific collision rules between packets. In each case, the same admissibility principle $\Delta D \leq 0$ governs updates; only motif statistics—and thus event propensities—change.

Thus, dense packings instantiate a general recipe for computation from a distribution: (1) declare τvia observables; (2) choose a distribution over micro-features; (3) drive quasi-statically so events are clean jumps; (4) read out computation from flip fields, defect kinematics, and D's staircase. The "program" is the distribution; the "processor" is mechanics; the "correctness" certificate is monotone descent of $D(q \parallel \tau)$.

2. Framework

2.1 Definitions: q, τ, observable set, $D(q \parallel \tau)$

State q.
The realized microstate: particle positions/orientations and the derived contact/neighbor topology, defect inventory, spectra, etc. We never compare raw coordinates; we compare observables derived from them.

Observable set $\mathcal{O}$.
A declared, finite set of statistics extracted from q. Typical choices (pick what the paper needs):

- Contact/force: mean contact number Z, residual overlap energy E_{ol}, isostaticity gap $Z - Z_{\text{iso}}$.

- Topology/defects: 5–7 Voronoi defect densities, dislocation/disclination counts, loop/cluster motifs in the contact graph.
- Orientational/positional order: | ψ_6 |distribution, bond-orientational correlation length ξ_6, pair correlation $g(r)$.
- Spectral: structure factor $S(k)$ at small k (hyperuniformity diagnostics), radial spectra of the flip field.
- Dynamical/drive-based: event/collision rate under a specified cyclic drive (random organization).

Let $F(q) \in \mathbb{R}^m$ denote the vector of these observables (histograms may be embedded via sufficient coordinates).

Target τ.
A constraint-defined manifold (or distribution over $\mathcal{O}$) specifying what "aligned" looks like—for example:

- Isostatic τ: $Z = Z_{\text{iso}}, E_{\text{ol}} = 0$.
- Hexatic τ: target histogram of | ψ_6 |and reduced free-disclination density.
- Hyperuniform τ: specified small-k scaling of $S(k)$.
- Absorbing τ: zero future-event rate under the chosen drive.

Concretely, τ is represented by a reference vector $F(\tau)$ or a reference distribution $P_\tau(\mathcal{O})$.

Misfit $D(q \parallel \tau)$.
A nonnegative divergence measuring how far $F(q)$is from τ, e.g.:

- KL on histograms: $D_{\text{KL}}(P_q \parallel P_\tau) = \sum_i \quad P_q(i)\log \frac{P_q(i)}{P_\tau(i)}$.
- Quadratic form / Mahalanobis: $D = \frac{1}{2}(F(q) - F(\tau))^\top W \big(F(q) - F(\tau)\big)$ with positive semidefinite W.
- Kernel/Wasserstein options for spectral targets (e.g., MMD on $S(k)$, W_2 on $g(r)$).

Requirements: $D \geq 0$, $D = 0 \iff F(q) \in \tau$; stable under coarse-graining of $\mathcal{O}$; computable along a trajectory.

2.2 Process principle: one path, multiple clocks (time, iteration)

We distinguish the path from the clocks.

- Path Γ. The sequence of microstates $\{q\}$generated by the (fixed) physics plus protocol, projected to $\mathcal{O}$. Think of $\Gamma = \{F(q_s): s \in [0,1]\}$, parameterized by an abstract arclength $s$along the alignment route toward τ.
- Clocks.
    - Physical time t: resolves elastic waves, dissipation, contact reforming.
    - Iteration n: protocol ticks (compression increments, cycle counts).
    - Arclength s: intrinsic progress on Γ (clock-free progress).

Near jamming under quasi-static drive, events exhibit a timescale separation

$$\tau_{\text{drive}} \gg \tau_{\text{event}},$$

so D shows plateaus punctuated by jumps: piecewise-constant in n with discrete drops at event indices. Thus, "instantaneous tunneling" is instantaneous in iteration n (one admissible update), not in physical t.

A smooth idealization of the same idea is the alignment flow

$$\dot{q} = -\mathcal{G}^{-1}\nabla_q D(q \parallel \tau), \dot{D} \leq 0,$$

with $\mathcal{G}$ an appropriate (possibly state-dependent) metric; discrete events are then jump discontinuities along the same Γ.

2.3 Admissible descent: $\Delta D \leq 0$ on declared observables

Admissible update (definition). A state transition $q \to q'$ (e.g., a T1 flip, a plastic micro-slip, or a full avalanche treated as one compound step) is admissible iff

$$\Delta D \equiv D(q' \| \tau) - D(q \| \tau) \leq 0$$

with respect to the declared observable set $\mathcal{O}$. This is our only normative criterion.

Consequences.

- Monotonicity (staircase law): Along any trajectory built from admissible updates, D is nonincreasing; under quasi-static drive it exhibits plateaus and drops.
- Clock invariance: Whether we index by t or n, admissibility is judged only by ΔD on $\mathcal{O}$.
- Stationary/absorbing sets: $q^\star$ is stationary if no admissible local update exists ($\nabla_q D = 0$ for smooth flows); under cyclic drive, absorbing τcorresponds to vanishing event rate with D minimal on $\mathcal{O}$.
- Composability: A cluster of local events is admissible as a block if the net $\Delta D \leq 0$, allowing us to treat avalanches as single tunneling steps on Γ.
- Programmability by distribution: Physics and admissibility are fixed; changing only the distribution (sizes/shapes) changes the *propensities* of local motifs to trigger admissible steps, thereby altering the induced "rule table" without writing rules.

Implementation notes (for experiments/simulations).

1. Declare $\mathcal{O}$ for the chosen τ(e.g., | ψ_6 | histogram + defect density, or small-k $S(k)$ bins).
2. Choose D (KL on those histograms, quadratic form on moments, MMD on spectra).
3. Drive quasi-statically (compression steps or weak cyclic forcing) with deterministic relaxation.
4. Log events and compute ΔDper tick; visualize flip-field rasters and the staircase $D(n)$.
5. Identify stationary sets by absence of admissible local moves (no contact changes/collisions under the drive).

3. Local Events as Tunneling

3.1 T1/STZ as single-site barrier crossings

Phenomenology.

- T1 (2D foams/packings): a four-body neighbor exchange (one Delaunay edge disappears, the transverse edge appears).
- STZ (amorphous solids): a small cluster executes a shear-like flip between two orientational variants.

q→τ view.

- Local state q_{loc}: positions/orientations of the 3–6 nearest neighbors, the local contact graph, and residual overlap/stress.
- Target τ_{loc}: an admissible, lower-misfit contact topology (e.g., more isostatic, lower local residual).
- Tunneling step: a discrete map $q_{\mathrm{loc}} \to q'_{\mathrm{loc}}$ with $\Delta D_{\mathrm{loc}} \leq 0$ that passes a geometric barrier (edge flip or variant switch).

Metrics & diagnostics.

- Local misfit: $D_{\mathrm{loc}} = \alpha E_{\mathrm{ol}} + \beta(Z - Z_{\mathrm{iso}})^2 + \gamma\, \mathrm{KL}(|\ \psi_6\ |_{\mathrm{loc}} \| \ \ |\ \psi_6\ |_{\tau})$.
- Trigger condition: sign change of the unstable eigenmode of the local Hessian / contact Laplacian; or a geometric predicate (circumcircle swap in Delaunay).
- Signature: sharp drop in D_{loc}; non-affine displacement burst localized to 3–8 particles; update is instantaneous in iteration n.

Predictions.

- Probability a specific local motif tunnels on the next tick is a function of distribution-driven parameters (e.g., size triplet (r_i, r_j, r_k)).
- Post-hoc classifier "flip next tick?" trained on local motifs achieves high accuracy → the induced rule table.

Protocol.
Detect candidate edges (T1-ready quads), compute D_{loc} pre/post flip, log $\Delta D_{\mathrm{loc}} \leq 0$, and rasterize flips as a binary flip field.

3.2 Avalanches as multi-site tunneling; staircase $D(t)$

Phenomenology.
A T1/STZ can destabilize neighbors via elastic/constraint couplings, triggering cascades that span many sites before the system re-stabilizes.

q→τ view.

- Compound step: an avalanche is a *block update* $\mathcal{A} = \{q^{(0)} \to q^{(1)} \to \cdots \to q^{(k)}\}$ treated as one tunneling move along the same path Γ.

- Admissibility: $\Delta D_{\mathcal{A}} = D(q^{(k)} \parallel \tau) - D(q^{(0)} \parallel \tau) \leq 0$, regardless of the internal micro-steps.

Global signatures.

- Staircase law: $D(t)$ (or $D(n)$) shows plateaus separated by drops of size ΔD; between drops, $\dot{D} \approx 0$.
- Size statistics: avalanche extent S (flips or affected particles) and $|\Delta D|$often exhibit broad/heavy-tailed distributions; $|\Delta D|$correlates with S.
- Information flow: increased transfer entropy along the contact graph during cascades; defect packets propagate ballistically or sub-ballistically depending on motif density.

Measurables.

- Flip-field connected components per tick → avalanche size S.
- ΔD per avalanche (KL on $|\psi_6|$, on motif histograms, or small-k $S(k)$).
- Spatial kernels: impulse response—fit decay of flip probability with graph distance from the seed.

Distribution control (computation from a distribution).

- Narrow monodisperse: small, rare avalanches; fast healing to crystal-like τ.
- Tuned bimodal / shape-doped: frequent traveling packets; larger $|\Delta D|$; richer collision outcomes (annihilation, reflection, binding).

Protocol.
Quasi-static drive with “event-driven” relaxation (complete all flips before next tick); segment $D(t)$ to extract plateau/drop structure; report $(S, \Delta D)$ joint statistics.

3.3 Absorbing τ under cyclic drive (random organization)

Setup.
Apply a small-amplitude, fixed periodic protocol (e.g., sinusoidal shear or oscillatory compression) and observe whether collisions/contact changes persist.

q→τ view.

- Observable set $\mathcal{O}$: per-cycle event rate R, flip count, return distance to initial configuration.

- Absorbing τ: the subset where future cycles generate no new events (e.g., $R = 0$), i.e., the system is dynamically compatible with the drive.

Tunneling interpretation.
Each cycle enacts an admissible descent step if it reduces $D(q \parallel \tau)$ defined on $\mathcal{O}$ (e.g., KL between current and zero-event distribution). The system plateaus at nonzero R, then tunnels via an organizing avalanche to $R = 0$ (or to a minimally active steady state near threshold).

Diagnostics.

- Event-rate decay $R(n)$: exponential/critical slowing near the threshold amplitude; on-τ implies $R \to 0$.
- Cycle-to-cycle KL: $D_{\mathrm{KL}}\big(P_q(R) \parallel \delta_0\big)$ decreases in steps.
- Return map distance over cycles → convergence to a limit cycle (absorbing set) or to a finite-activity attractor above threshold.

Predictions & controls.

- Threshold behavior: below a distribution-dependent amplitude, the ensemble flows to absorbing τ; above it, sustained activity persists.
- Distribution tuning: size/shape heterogeneity shifts the threshold and alters the pathway (number/size of staircase drops).

Protocol.
Fix amplitude; run many cycles with full relaxation each half-cycle; log $R, \Delta D$ per cycle; identify convergence to absorbing τ(no events for Kconsecutive cycles) and map threshold vs distribution parameters.

Unifying note.
Sections 3.1–3.3 are the same story at three scales: single-site tunneling (T1/STZ), multi-site tunneling (avalanches, staircase D), and global tunneling to absorbing τunder cyclic drive. In all cases, admissible descent ($\Delta D \leq 0$ on declared observables) and distribution-shaped propensities replace explicit rules, delivering computation from a distribution along one alignment path $q \to \tau$.

4. Targets (τ) & Metrics (D)

4.1 Isostatic closures and contact number

Target (τ_iso). Frictionless disks in 2D are isostatic at jamming with mean contact number

$$Z_{\text{iso}} = 2d = 4 \text{(rattlers excluded)}.$$

Locally, force/torque balance holds and residual overlaps vanish.

Observable set $\mathcal{O}_{\text{iso}}$.

- Global: $\bar{Z}$ (rattler-trimmed), contact-number histogram $P_q(Z)$, residual-overlap energy E_{ol}, force-balance residual R_f.
- Local: distribution of coordination per particle, per-species Z in mixtures, fraction of rattlers ρ_{rat}.

Metrics $D(q \parallel \tau_{\text{iso}})$.

- Quadratic: $D_Z = \frac{1}{2} w_Z (\bar{Z} - 4)^2 + \frac{1}{2} w_{\text{rat}} \rho_{\text{rat}}^2 + w_E E_{\text{ol}}$.
- Histogram-KL: $D_{\text{KL}}\left(P_q(Z) \parallel \delta_{Z=4}\right)$ (practically, a narrow reference around 4).
- Constraint residual: Mahalanobis norm of force-balance residuals.

Admissible update. Any micro-rearrangement with $\Delta D_Z \leq 0$: closes missing contacts, removes overlaps, or demotes rattlers by creating stable contacts.

Protocol notes. Trim rattlers consistently; report both $\bar{Z}$and $P_q(Z)$; for finite systems, compare to a smoothed reference $P_\tau(Z)$ centered at 4.

4.2 Hexatic/defect statistics (ψ_6, 5–7 pairs)

Target (τ_{\hex}). High bond-orientational order with suppressed free disclinations; for perfect hexagonal order, | ψ_6 |= 1and isolated 5/7 defects are absent.

Observables $\mathcal{O}_{\backslash \text{hex}}$.

- $\psi_6(i) = \frac{1}{k_i} \sum_{j \in \mathcal{N}_i} \quad e^{i 6 \theta_{ij}}$; collect the histogram H_q(| ψ_6 |).
- Correlation length $\xi_6$from $C_6(r) = \langle \psi_6(0) \psi_6^{\backslash *}(r) \rangle$.

- Voronoi coordination: densities ρ_5, ρ_7 and bound-pair vs free-disclination counts; defect-cluster size distribution.

Metrics $D(q \parallel \tau_{\backslash \text{hex}})$.

- KL on order parameter: $D_\psi = D_{\text{KL}}(H_q(|\psi_6|) \parallel H_\tau(|\psi_6|))$ (take H_τ from a hexatic/hex crystal reference at matching density).
- Defect penalty: $D_{\text{def}} = w_5\rho_5 + w_7\rho_7 + w_b$ (free disclinations).
- Correlation gap: $D_{C_6} = \sum_r \quad w(r)\,[C_{6,\tau}(r) - C_{6,q}(r)]^2$.

Admissible update. A T1/STZ that binds a 5–7 pair, annihilates free disclinations, or increases $|\psi_6|/\xi_6$ so that $\Delta(D_\psi + D_{\text{def}} + D_{C_6}) \le 0$.

Protocol notes. Use the same neighbor definition (Voronoi or fixed-k) across runs; for mixtures, stratify H_q by species or weight by area.

4.3 Spectral targets $S(k)$ (hyperuniformity)

Target (τ_{\HU}). Hyperuniform spectra suppress long-wavelength number fluctuations, typically $S(k) \to 0$ as $k \to 0$; for MRJ disks in 2D, empirical small-k scaling is close to linear $S(k) \sim Ak$.

Observables $\mathcal{O}_{\backslash \text{HU}}$.

- Angle-averaged structure factor $S(k)$binned over $k \in [k_{\min}, k_c]$ (with careful windowing/PBC).
- Real-space number variance $\sigma_N^2(R)$ in sampling windows of radius R (for cross-checks).

Metrics $D(q \parallel \tau_{\backslash \text{HU}})$.

- Small-k spectral loss: $D_S = \sum_{k \le k_c} \quad w_k\,[S_q(k) - S_\tau(k)]^2$ or KL on normalized binned spectra.
- Slope misfit: $D_{\text{slope}} = (\hat{A}_q - \hat{A}_\tau)^2$ for a linear fit $S(k) \approx \hat{A}\,k$.
- Number-variance gap: $D_\sigma = \sum_R \quad w(R)\,[\sigma^2_{N,q}(R) - \sigma^2_{N,\tau}(R)]^2$.

Admissible update. Any rearrangement whose net effect reduces long-wavelength power (e.g., compaction that smooths large-scale density fluctuations): $\Delta D_S \le 0$.

Protocol notes. Use identical box sizes/PBC; fix $k_c$below first Bragg-like features; bootstrap confidence on $\hat{A}$. Report both spectral and real-space diagnostics for robustness.

4.4 Motif libraries; KL on motif frequencies

Target (τ_{\motif}). A library of favorable local and mesoscale motifs with target frequencies—e.g., contact-graph patterns, Delaunay/Voronoi tiles, size/shape triplets, junctions seeded by dopants.

Motif construction.

- Local contact motifs: unlabeled/labeled graphlets on 3–6 nodes (e.g., triangles, rhombi, 4-cycles with diagonals), optionally augmented with species/size bins.
- Geometric tiles: Delaunay triangle types (acute/obtuse classes), Voronoi polygon types.
- Heterogeneity motifs: (radius or shape) neighborhoods—e.g., (r_i, r_j, r_k) binned ratios; disk–triangle interlocks; dopant-centered junctions (Y/T/χ).

Build a dictionary $\mathcal{M} = \{m_1, \dots, m_M\}$. Let $f_q(m)$ be the observed frequency (normalized) and $f_\tau(m)$ the target (from a designed τ or from a high-performing reference ensemble).

Metrics $D(q \parallel \tau_{\backslash\text{motif}})$.

- Motif-KL: $D_{\mathcal{M}} = D_{\mathrm{KL}}(f_q \parallel f_\tau) = \sum_{m\in\mathcal{M}} f_q(m) \log \frac{f_q(m)}{f_\tau(m)}$ (use Laplace smoothing).
- Weighted gap: $D^{\wedge}_{\mathcal{M}} = \sum_m w_m [f_q(m) - f_\tau(m)]^2$ (emphasize "functional" motifs, e.g., junctions that launch traveling defects).
- Predictivity loss: train a deterministic classifier $h$that predicts "flip next tick?" from the local motif; define $D_h = 1 - \mathrm{AUC}(h)$. Lower D_h implies a crisper induced rule table.

Admissible update. Any rearrangement that moves motif frequencies toward f_τ: $\Delta D_{\mathcal{M}} \leq 0$. In practice, a T1 that converts an overrepresented "frustrated" motif into a target motif is counted as tunneling.

Protocol notes.

- Choose $\mathcal{M}$at the resolution that balances parsimony (few bins) with discriminability (separates active vs inactive neighborhoods).
- For mixtures, include species tags but keep bins coarse (e.g., small/large) to maintain sample efficiency.
- Validate with transfer entropy: motifs that reduce $D_{\mathcal{M}}$should also increase directed information flow along plausible propagation directions.

Putting it together

For any declared τ, set $\mathcal{O}$ and an accompanying D (or a weighted sum $D = \sum_i \lambda_i D_i$). A tunneling step—single T1/STZ, compound avalanche, or one organizing cycle—must satisfy $\Delta D \leq 0$. The same physics and admissibility criterion apply across targets; only the observables and motif priors change, which is why *programming by distribution* (rather than by explicit rules) cleanly fits the q→τ framework.

5. Induced Local Logic from a Distribution

5.1 Uniform disks; doped shapes/size mixtures

Premise. We fix physics (hard non-overlap + deterministic relaxation), protocol (quasi-static steps or weak cyclic drive), and admissibility ($\Delta D \leq 0$). The only knob is the distribution $\mathcal{D}$over microscopic attributes (sizes/shapes/orientations). Changing $\mathcal{D}$ alters which local motifs are frequent and mechanically precarious, thereby changing the probabilities that specific neighbor exchanges (T1/STZ) become admissible at the next tick. That is the induced local logic.

A. Uniform disks (control).
$\mathcal{D}$: $r \equiv r_0$. With PBC and gentle densification, the system selects the hexagonal τ; defects are rare and quickly annihilate. Diagnostics: rapid fall of $D_{\backslash \text{hex}}$; avalanche size Snarrowly distributed; flip-field entropy rate low.

B. Size mixtures (bidisperse/polydisperse).
$\mathcal{D}$: $r \sim (1-\phi_\ell)\delta(r-r_s) + \phi_\ell \delta(r-r_\ell)$ with ratio $\rho = r_\ell / r_s$ and large-fraction ϕ_ℓ.

- Frustration window: $\rho \approx 1.2 - 1.35$, $\phi_\ell \in [0.1, 0.35]$ tends to suppress perfect $\tau_{\backslash \text{hex}}$ yet sustain mobile defect strings.
- Effect: abundant T1 seeds at small–small–large triplets; traveling defect "packets" emerge; avalanches broaden; $D_{\mathcal{M}}$ (motif-KL) drops in steps.

C. Shape doping (e.g., equilateral triangles among disks, equal area).
$\mathcal{D}$: with probability $\phi_\triangle$ place a triangle (random orientation), else a disk.

- Triangles create junction motifs (Y/T/χ) that template branching and reflection of defect packets.
- At low $\phi_\triangle$(<10%) they act as rare scatterers; at moderate $\phi_\triangle$ (15–30%) they form a sparse routing network for packets, altering collision logic without external rules.

Targets (τ) under these $\mathcal{D}$.

- τ $_{\backslash hex}$ (hexatic), τ $_{\backslash HU}$ (hyperuniform spectral), τ $_{abs}$ (absorbing under a fixed cyclic amplitude), τ $_{\mathcal{M}}$ (motif library). The same $\mathcal{D}$ can be assessed against different τ by swapping $\mathcal{O}$ and D.

5.2 Post-hoc "rule tables" via neighborhood motifs

Goal. Extract the induced rule "flip next tick?" from motifs determined by $\mathcal{D}$, without writing rules a priori.

Neighborhood representation.
For each candidate Delaunay edge $e = (i, j)$ (or local 4-body quad), define a feature map $\Phi_e$that is purely geometric/combinatorial, e.g.:

- Topological: local contact subgraph G_e on nodes within two hops; degree labels; cycle counts.
- Geometric: sorted radii in $\mathcal{N}_e$ (ratios $\{r_k/r_i\}$ binned), edge length ratios, angles.
- Shape flags: presence of triangle(s), triangle–disk interlock types.
- Stress proxies: local overlap residuals or force imbalances if a penalty solver is used.

Motif library $\mathcal{M}$.
Discretize Φ_e to a finite set of motif types $\{m\}$. For each tick, label $y_e \in \{0,1\}$ ("edge $e$flips next tick?"). Estimate:

$$p_{\mathcal{D}}(y = 1 \mid m) = \frac{\#\text{ of flips with motif } m}{\#\text{ of trials with motif } m}.$$

This conditional is the induced rule table under $\mathcal{D}$.

Quality measures.

- Predictivity: train a simple deterministic classifier $h(\Phi_e)$ (e.g., decision tree / logical rules) on one run; report AUC/accuracy on another run with the same $\mathcal{D}$. High AUC ⇒ crisp induced rules.
- Stability: compare $p_{\mathcal{D}}(\, y = 1 \mid m\,)$ across seeds; compute variation distance to confirm distribution-level invariance.
- Causality (direction): compute transfer entropy $TE(m_t \to y_{t+1})$ along graph directions; significant TE for specific motifs indicates signal-carrying capacity.

Tunneling consistency.
Motifs with large $p_{\mathcal{D}}(y = 1 \mid m)$ should correlate with negative ΔD_{loc} (admissible descent). We therefore log ΔD_{loc} per flip; a compact rule table that reliably predicts $\Delta D_{\text{loc}} < 0$ validates the q→τ tunneling interpretation.

5.3 Phase diagram over moments of the distribution

Axes. Use moments of $\mathcal{D}$ as macroscopic coordinates:

- Size mixtures: mean fixed; $m_2 = \mathrm{Var}(r)/\langle r \rangle^2$(polydispersity), bimodality via $\rho = r_\ell / r_s$and ϕ_ℓ.
- Shape doping: $\phi_\triangle$ (shape fraction), optional orientation variance $m_{2,\theta}$.
- General: skewness m_3, kurtosis m_4 if needed.

Regimes (qualitative).

1. Crystal-seeking: small m_2, $\phi_\triangle \approx 0$. Rapid approach to τ $_{\text{\hex}}$; small avalanches; low entropy rate; $D_{\text{\hex}} \to 0$.
2. Automaton-like: intermediate $m_2$or $1.2 \lesssim \rho \lesssim 1.35$ with $\phi_\ell \in [0.1, 0.35]$; or $\phi_\triangle \in [0.15, 0.30]$. Mobile, interacting defect packets; broad avalanche sizes; staircase $D(t)$; high motif predictivity h; robust TE along packet paths.
3. Glassy/jammed: large m_2 or poorly tuned mixtures; frozen heterogeneity; many micro-events but little propagation; Ddecays slowly; flip-field entropy high but TE low (noise-like).

Order parameters for the diagram.

- Rule sharpness: $R_h = 1 - \mathrm{AUC}(h)$ (lower is better "rule").
- Packet mobility: mean packet speed v_p from defect tracking; fraction of flips in propagating clusters.
- Avalanche statistics: power-law tail exponent τ_S (if present), cutoff S_c.
- Information flow: average TE along high-degree edges; entropy rate H of the flip field.
- Alignment: endpoint $D$to chosen τ (e.g., $D_{\text{\HU}}$at small k; $D_{\text{\hex}}$).

Protocol to map the diagram.

1. Choose τ and $\mathcal{O}$ (e.g., τ \hex + motif library).
2. Sweep $\mathcal{D}$over a grid in $(m_2, \rho, \phi_\ell, \phi_\triangle)$.
3. For each point: run quasi-static PBC simulation (deterministic solver), compute $D(t)$, extract flips/avalanches, build motif tables, fit h, compute R_h, v_p, τ_S, TE, H.
4. Color the diagram by automaton index

$$\mathcal{A} = w_1(1 - R_h) + w_2\,\hat{v}_p + w_3\,\mathrm{TE} - w_4\,H,$$

normalize to $[0,1]$. High $\mathcal{A}$ marks the CA-like regime.

Predictions.

- There exists a ridge in (ρ, ϕ_ℓ) where $\mathcal{A}$ peaks (maximal induced logic): too little heterogeneity ⇒ crystal; too much ⇒ glass.
- Shape doping shifts/steepens the ridge and introduces junction-logic (branching/reflection probabilities) visible in motif-KL trajectories.

Takeaway. A single, fixed mechanics+protocol+admissibility stack yields programmable emergent logic: the program is $\mathcal{D}$. By tuning moments (variance, bimodality, doping fraction), we move between crystal-seeking, automaton-like, and glassy regimes, and we can read off the induced rule table and tunneling staircases entirely from observable descent $D(q \parallel \tau)$.

6. Protocols & Diagnostics

6.1 Quasi-static compression / weak periodic drive with PBC

Geometry & boundaries.

- 2D square torus (periodic boundary conditions, PBC), box vectors $L_x = L_y = L$.
- Nondimensionalize so mean disk diameter $= 1$; report volume fraction ϕ.

Initialization.

- Place $N \in [2 \times 10^3, 2 \times 10^4]$ particles by Poisson-disk or blue-noise.
- Draw sizes/shapes from the chosen distribution $\mathcal{D}$ (uniform disks; bidisperse; doped triangles, etc.).

- Minimize a soft overlap energy (e.g., quadratic penalty) to remove gross overlaps at low ϕ.

Deterministic solver (event-driven quasi-static).

- Compression mode: increase a scalar control c (e.g., particle radii or uniform box shrink with PBC) in tiny steps Δcuntil first overlap/contact-graph change; freeze $c$and minimize until no overlaps remain (up to tolerance). That "relaxation burst" is one event block. Repeat.
- Periodic mode (weak drive): impose an oscillatory affine map (shear or dilation) $A(t) = A_0 \sin\left(\frac{2\pi t}{T}\right)$ with small amplitude. For each half-cycle: drive → relax to force/overlap tolerance → log events.

Recommended numerics.

- Penalty potential $\phi(\delta) = \frac{1}{2} k\, \delta^2$ for $\delta > 0$(overlap), else 0.
- CG/L-BFGS tolerance on force norm $< 10^{-10}$; overlap tolerance $< 10^{-12}$.
- Quasi-static step size: adapt Δcso that each step triggers 0 or 1topological change before relaxation; shrink Δcon multiple triggers.
- Randomness: fix seeds; make all tie-breaks deterministic (lexicographic ordering of edges, stable sorting of neighbors).

Outputs per event block (tick n).

- Contact graph G_n (with species/size tags).
- Set of changed edges $\mathcal{E}_n$ (creations/deletions).
- Particle displacements u_n; non-affine measure $D_{\min}$.
- Observables $F(q_n)$ (contact number, $|\ \psi_6\ |$histogram, defect counts, $S(k)$, etc.).
- Misfit(s) $D(q_n \parallel \tau)$.

6.2 Flip-field rasters, defect tracking, transfer entropy

Flip-field raster.

- Define a binary field on particles (or edges):

$$X_i(n) = \begin{cases} 1 & \text{if particle } i \text{ participates in any edge flip in block } n \\ 0 & \text{otherwise.} \end{cases}$$

- Visualize as a time–index image (rows = ticks n, columns = particle IDs sorted by a spatial key, e.g., Hilbert/PBC order).
- Avalanche segmentation: Connected components in the bipartite graph (time adjacency + spatial contact adjacency) within tick $n$give avalanche size S_n.
- Entropy rate: estimate $H\big(\, X(n) \mid X(n-1) \,\big)$ from empirical frequencies (add Laplace smoothing).

Defect tracking (Voronoi 5–7).

- Construct PBC Voronoi; label each cell by coordination z.
- Defects: 5- and 7-coordinated sites; dislocations are bound $5 - 7$ pairs (within a cutoff).
- Track defect cores across ticks by nearest-neighbor matching under PBC with gating on displacement.
- Kinematics: defect velocity v, mean free path, annihilation/creation events; packet = coherent cluster of defects moving together.
- Order metrics: ρ_5, ρ_7; correlation length of ψ_6; reduction in free disclinations signals admissible descent for hexatic τ.

Transfer entropy (TE) on the contact graph.

- Build a directed, time-lagged graph: for each edge $(i \to j)$ if $i$and j are neighbors at tick n.
- Variables: $X_i(n) \in \{0,1\}$ (flip participation).
- Estimate pairwise TE at lag 1:

$$\mathrm{TE}_{i\to j} = \sum_{x_j, x_j', x_i} p(x_j', x_j, x_i) \log \frac{p(x_j' \mid x_j, x_i)}{p(x_j' \mid x_j)}.$$

- Use plug-in estimator with Miller–Madow correction; assess significance by time-shuffling $X_i$to form a null TE $_0$; report $\Delta\mathrm{TE} = \mathrm{TE} - \mathrm{TE}_0$.
- Diagnostics: (i) mean TE along presumed propagation directions, (ii) TE anisotropy (fronts), (iii) TE concentration on packet backbones.

Local motif mining (for induced rules).

- For each candidate edge, encode a motif from its radius/shape neighborhood and topological pattern.
- Estimate $p(\,y = 1 \mid m\,)$ where y= "flip next tick."
- Train a deterministic, sparse predictor (e.g., depth-3 decision tree); report AUC on held-out seeds; rule sharpness $= 1 -$ AUC.

6.3 Reproducible staircases; hysteresis tests

Staircase extraction (monotone D).

- Compute $D(q_n \parallel \tau)$each tick; apply a small total-variation denoiser (optional) to remove numerical jitter while preserving drops.
- Identify plateaus where $| D_{n+1} - D_n | < \epsilon$ (e.g., 10^{-6}) for $\geq m$ ticks, and jumps where $\Delta D_n = D_{n+1} - D_n \ll -\epsilon$.
- Report distribution of jump sizes $|\Delta D|$, residence times on plateaus, and correlation with avalanche size S.

Reproducibility.

- Fix $\mathcal{D}$, protocol, and seed; run Rreplicates with microscopic perturbations (e.g., permuted particle indices but same $\mathcal{D}$).
- Quantify variability of $D(t)$ staircases via dynamic time warping (DTW) distance; align jumps; compute mean/joint CIs by bootstrap.
- Stability of induced rules: compare motif tables $p(\,y = 1 \mid m\,)$ across replicates (total variation distance); high similarity ⇒ distribution-level law, not noise.

Hysteresis (path dependence).

- Compression–decompression loop: compress quasi-statically to $c_{\max}$, then decompress to starting c, logging D, S, defects.
    - Signature: irreversibility (non-closed Dloop), retained defect density; quantify loop area $\oint D\, dc$.
- Amplitude sweeps (periodic drive): ramp drive amplitude A up and down across the random-organization threshold; measure cycle event rate $R(A)$.

- Signature: absorbing τbelow threshold ($R \to 0$), finite-activity attractor above; look for hysteresis in $R(A)$ and in TE concentration.

Controls & robustness.

- PBC vs walls: confirm that qualitative results (staircases, packet kinematics, induced rules) persist under PBC; walls can seed artifacts.
- Tolerances: vary solver tolerances to ensure ΔD drops are not numerical.
- Finite-size scaling: repeat at $N \in \{2{,}048, 8{,}192, 32{,}768\}$; check scaling of jump sizes, TE peaks, and spectral misfit D_S at small k.
- Target swap: repeat diagnostics for multiple τ(isostatic, hexatic, hyperuniform, absorbing) to verify admissible descent is target-agnostic.

Reporting checklist (per figure/table).

- Specify $\mathcal{D}$ (moments, mixture ratios, shape fractions).
- Protocol (compression step rule or cyclic amplitude; PBC; solver tolerances).
- N, runtime ticks, random seed.
- Which τ, which $\mathcal{O}$, which D (exact formula and weights).
- Summary metrics: avalanche S, jump sizes | ΔD |, rule sharpness, TE, entropy rate, endpoint D.
- Reproducibility statistics and hysteresis loop area.

These procedures make the q→τ claims falsifiable: (i) admissible descent ($\Delta D \leq 0$) at event blocks, (ii) staircase $D(t)$ under quasi-static protocol, (iii) distribution-induced rules that predict flips, and (iv) approach to absorbing τunder weak periodic drive.

7. Parallels to Learning by Compression (LLMs) — comparative, not mechanistic

7.1 Mapping: q, τ, D $\leftrightarrow$ parameters, data/prior, loss/KL

Packing side (q→τ):

- State q: realized microstate summarized by observables (contacts, defects, spectra).
- Target τ: constraint-defined manifold (isostatic, hexatic/hyperuniform statistics, absorbing sets).

- Misfit $D(q \parallel \tau)$: divergence on declared observables (e.g., KL on | ψ_6 |histogram, small-k $S(k)$, motif frequencies).

LLM side (learning by compression):

- State (parameters w): current weights + induced representation.
- Target (data/prior): empirical data distribution $\mathcal{D}_{\text{data}}$+ inductive biases/regularizers (priors, model class).
- Misfit (loss/KL): $\mathcal{L}(w) \approx D_{\text{KL}}(p_{\text{data}} \parallel p_w)$ plus regularization; compression shows up via flat/minimal-description solutions.

Process mapping (operational):

- Update: admissible descent $\Delta D \leq 0 \leftrightarrow$ SGD step $\Delta \mathcal{L} \leq 0$.
- Clocks: iteration ticks / cycles $\leftrightarrow$ optimization steps / epochs.
- Stationarity: $\nabla_q D = 0 \leftrightarrow \nabla_w \mathcal{L} = 0$ (train loss flat; stable representation).

Read this as a shape-of-process analogy: both systems descend a misfit subject to constraints/priors; neither requires hand-coded local rules for emergent structure.

7.2 Plateau-and-jump descent; representation phase changes

Packings (q→τ).
Under quasi-static drive, $D(q \parallel \tau)$exhibits plateaus (metastable basins) and jumps (T1/STZ/avalanche tunneling). Defect packets rewire the contact graph; global observables (e.g., | ψ_6 |, $S(k)$) shift in discrete steps.

LLMs.
Training curves often show plateau-and-break dynamics: long stretches of slow loss decrease punctuated by representation phase changes (e.g., sudden emergence of syntax/algorithmic circuits, grokking transitions, sparsity onsets). Compression/flatness increases after these breaks (evidence of simpler internal codes).

Comparative diagnostics.

- Staircase curves: $D(n) \leftrightarrow$ loss/metric curves with elbow points.
- Event fields: flip rasters/TE peaks $\leftrightarrow$ activation-pattern change points, circuit saliency shifts.

- After-jump order: increased | ψ_6 | hyperuniformity ↔ sparser, more linearly separable features; flatter minima.

7.3 Emergent modularity ↔ hyperuniform/defect-templated order

Packings.

- Hyperuniform τ: low-k suppression indicates globally coordinated structure from local constraints.
- Defect templating: dopants (size/shape) seed junction motifs; defect packets follow reproducible routes—modular substructures in space.

LLMs.

- Modularity: tasks/skills localize into subnetworks (attention heads/MLP channels) with sparse inter-module traffic—an internal wiring diagram emerging from data statistics and optimization.
- Compression: implicit bias favors lower-description-length representations (weight sparsity, low-rank structure, flat minima).

Bridge ideas.

- Motif libraries ↔ circuits: motif frequency KL decreasing ↔ consolidation of functional subgraphs (circuits) consistent with data.
- Information routing: TE concentration along defect backbones ↔ increased information flow within task-relevant pathways (e.g., attention routes).
- Distribution programs: tuning $\mathcal{D}$(size/shape mix) to change induced logic ↔ changing data curriculum/augmentation to shape emergent circuits.

7.4 Caveats and limits of the analogy

- Mechanism mismatch. Packings obey geometric mechanics (non-overlap + relaxation); LLMs use backpropagation in high-dimensional parameter spaces. The analogy is process-level, not dynamical equivalence.
- Objective specification. In packings we declare observables and D; in LLMs, the objective is loss on data with implicit priors. Choice of $\mathcal{O}$ can shape "what counts" as descent.

- Clock separation. “Instantaneous” tunneling is instantaneous on the iteration clock, not in physical time; LLM “jumps” still take multiple optimization steps.
- Universality claims. We do not claim CA-universality or computational equivalence. We claim computation-from-a-distribution via admissible descent with similar staircase geometry and emergent modularity.
- Metric dependence. Conclusions depend on sensible D (KL on well-chosen observables). Poorly chosen metrics can hide tunneling in packings or phase changes in LLMs.

Use in paper. Present as a comparative sidebar with one figure: left—$D(n)$ staircase & flip raster; right—LLM loss with a known phase break & a schematic circuit emergence. Close with a one-paragraph caveat: same alignment geometry, different micromechanics.

8. Results Archetypes

8.1 Glider-like defect packets and collisions

What to see. In tuned mixtures (e.g., bidisperse with $\rho \approx 1.25$, $\phi_\ell \in [0.15,0.30]$ or modest shape-doping), localized strings of 5–7 Voronoi defects self-assemble into packets that move coherently across the contact network—our packing analog of CA “gliders.”

Quantification.

- Packet detection. Cluster co-moving 5–7 pairs by spatiotemporal connectivity; require minimum lifetime $T_{\min}$ and speed $v_p > 0$.
- Kinematics. Report distributions of v_p, lifetimes T, mean free paths ℓ, curvature of trajectories, and anisotropy under PBC.
- Information transport. Elevated transfer entropy (TE) along the packet backbone relative to null (time-shuffled) signals; TE peaks at the packet head, decaying behind.
- Tunneling signal. Each packet passage coincides with a local $\Delta D_{\mathrm{loc}} < 0$ (e.g., defect penalty and | ψ_6 |-KL drop), and aggregates into a global step in $D(n)$.

Collisions (logic-like interactions).

- Annihilation: head-on packets cancel, leaving a low-defect wake (net $\Delta D < 0$).
- Reflection: packet deflects from a dopant-templated junction (Y/T/χ motif), continuing with reduced speed.

- Binding: two packets merge into a slower composite traveling along a motif-preferred path.
  Report outcome probabilities conditioned on motif context; these are the induced local rules from the distribution, not hand-coded logic.

Figures. (i) Flip-field raster with packet world-lines; (ii) packet speed/lifetime histograms; (iii) TE map over the contact graph; (iv) $D(n)$ with collision-aligned drops.

8.2 Convergence to absorbing sets; random organization

Setup. Apply a low-amplitude periodic protocol (shear or dilation) with PBC. Below a distribution-dependent threshold, the system organizes into a state with no further contact changes per cycle—an absorbing τ.

Measurements.

- Event rate $R(n)$ (per cycle). Expect monotone decay, often with plateaus and discrete drops; convergence to $R \to 0$signals landing on absorbing τ.
- Cycle-KL. $D_{\mathrm{cyc}} = D_{\mathrm{KL}}\big(P_q(R) \parallel \delta_0\big)$ decreases in staircases; its final plateau defines the absorbing set.
- Return distance. Configurational distance to the cycle's start (invariant limit cycle → distance → 0).
- Basin mapping. Vary amplitude $A$to map the threshold $A^{\star}(\mathcal{D})$; below $A^{\star}$→ absorbing τ; above → finite-activity attractor (nonzero R).

Tunneling interpretation. Each "organizing burst" under the drive is a compound admissible step with $\Delta D < 0$, proceeding until $\nabla_q D = 0$ on the declared observables (no future events under the protocol).

Hysteresis & reproducibility.

- Hysteresis. Ramp A up/down; plot $R(A)$. Loop area reflects path dependence of organizing avalanches.
- Replicates. With fixed $\mathcal{D}$, absorbing microstates differ microscopically but share identical observable signatures (same final D, same motif frequencies within error), confirming a τ-manifold rather than a single point.

Figures. (i) $R(n)$ decay and staircase $D_{\mathrm{cyc}}(n)$; (ii) amplitude sweep $R(A)$ with threshold; (iii) return-map distance to the limit cycle.

8.3 Emergent hyperuniformity without explicit rules

Claim. Near jamming, distributions that promote mobile reorganization (but avoid deep glass) generate hyperuniform long-wavelength order (suppressed $S(k)$ as $k \to 0$) without any spectral objective in the dynamics—an emergent outcome of local admissible descent.

Measurements.

- Structure factor $S(k)$. Angle-average under PBC; fit small-k window $k \in [2\pi/L, k_c]$. Extract slope $\widehat{A}$ for linear behavior $S(k) \approx \hat{A}\, k$(2D MRJ-like).
- Spectral misfit. $D_S = \sum_{k \le k_c} w_k\, [S_q(k) - S_\tau(k)]^2$ or KL on binned spectra to a hyperuniform reference $S_\tau(k)$.
- Number variance. $\sigma_N^2(R)$ scaling with window radius R; reduced growth vs a Poisson or non-hyperuniform control.

Tunneling correlation. Large staircase drops in $D(n)$ align with reductions in small-k power and with decreases in defect density—i.e., cascades act as spectral "smoothers." Report cross-correlations between ΔD, $\Delta \hat{A}$, and avalanche size S.

Distribution dependence.

- Monodisperse control: fast approach to crystalline Bragg peaks (not the hyperuniform disordered target).
- Automaton-like window: intermediate variance/bimodality or moderate shape-doping → strongest suppression at small k while retaining disorder at higher k.
- Glassy regime: slow spectral improvement; residual small-k power persists.

Figures. (i) $S(k)$ curves across the run with small-k fits; (ii) spectral misfit $D_S(n)$ vs global $D(n)$; (iii) variance $\sigma_N^2(R)$ scaling; (iv) joint scatter $(S, \Delta\hat{A}, |\ \Delta D\ |)$.

Unifying note. Across 8.1–8.3, the signatures—glider-like packets, absorbing sets, hyperuniform spectra—all manifest as admissible descent in $D(q \parallel \tau)$: local flips tunnel to lower-misfit neighborhoods; cascades stitch many such tunnels into global steps; periodic driving guides the system to τ-manifolds where further events vanish. No explicit rules are written; the distribution selects the motifs and, with mechanics, induces the effective local logic.

9. Discussion & Outlook

9.1 Consistency with prior q→τ work

Our account keeps faith with the core tenets established across the q→τ corpus:

- One process, many clocks. We separate the *path* (alignment curve Γ) from *clocks* (physical time, iteration count, arclength), so "instantaneous tunneling" means *one admissible update on the protocol clock*, not literally zero physical duration. This is the same Path-not-Clock stance used in earlier alignment pieces.
- Declared observables determine the game. We never claim a unique τ; we *declare* an observable set $\mathcal{O}$and induce the corresponding $\tau(\mathcal{O})$. The same micro-trajectory can be descending for hexatic τ and *not* for a spectral τ, depending on $\mathcal{O}$. That mirrors prior work where different tasks/targets induce different misfits.
- Admissible descent as the only norm. Updates are "good" iff $\Delta D(q \parallel \tau) \leq 0$ on the declared observables. No extra teleology, no hidden rules.
- Tunneling parameterization. Our staircase drops are discrete realizations of the same informational tunneling path q_t, now seen in mechanical systems: plateaus (stuck in a basin) punctuated by jumps (admissible barrier crossings).
- Computation from a distribution. The *program* is the heterogeneity measure (sizes/shapes); mechanics + admissibility are fixed. This recapitulates earlier "law of becoming" arguments where changing the prior or data distribution changes the induced logic without altering the update machinery.

Together, these keep the packing narrative inside the same q→τ geometry—no new metaphysics, only a new *domain* where the geometry is testable with pixels and forces.

9.2 Open problems (proofs of monotonicity, optimal observables, 3D/finite-T)

(i) Monotonicity with discrete events.
We observe $\Delta D \leq 0$ at event blocks under quasi-static protocols, but a *proof* requires conditions on: (a) metric choice, (b) solver tolerances, (c) event segmentation (composing many micro-flips into one block). Two routes:

- Local sufficient conditions: show that each admissible T1/STZ strictly decreases a convex surrogate of D (e.g., a quadratic on a small | ψ_6 |moment set) when the local unstable mode crosses zero.

- Block monotonicity: prove that the relaxed post-event state is a minimizer of a convexified D + penalty functional in an event cone, hence D cannot increase across the block.

(ii) Optimal observables $\mathcal{O}^{\star}$.
KL on | ψ_6 |, motif frequencies, and small-k $S(k)$ all work—but are they *minimal and sufficient*? Open:

- Find low-dimensional summaries that preserve the staircase geometry.
- Characterize target identifiability: when do two different τproduce indistinguishable D traces?
- Learn $\mathcal{O}^{\star}$from data by maximizing (jump detectability, TE concentration, prediction of absorbing thresholds) under complexity penalties.

(iii) 3D, friction, and finite temperature.

- 3D spheres/polyhedra: T1 analogs become face/edge swaps; defect taxonomy changes; isostatic counts differ. Extend τ $_{iso}$, τ $_{\hex}$ (now crystalline/icosahedral order parameters), and τ $_{\HU}$ to 3D and re-test monotonicity.
- Friction & cohesion: Coulomb cones add non-convexity; contact creation/destruction is history-dependent. Need admissibility definitions that incorporate inequality constraints and stick–slip logic while retaining $\Delta D \leq 0$ in aggregate.
- Finite-T & noise: With thermal agitation or agitation noise, events blur. Replace $\Delta D \leq 0$almost surely with $\mathbb{E}[\Delta D] \leq 0$ and bound variance; test fluctuation–dissipation style relations for staircases.

(iv) Universality & criticality.
Do avalanche exponents and TE patterns fall into known universality classes (e.g., absorbing-state, depinning)? Can the automaton index (from Sec. 5.3) be shown to peak at a bona fide critical manifold in moment-space of the distribution?

(v) Inverse design (τ→$\mathcal{D}$).
Given a desired induced rule table or packet logic (e.g., reflect at junctions with prescribed probabilities), solve for a distribution $\mathcal{D}$(size ratios, dopant fractions, shapes) that maximizes rule sharpness subject to manufacturability.

(vi) Formal ties to learning theory.
Make the comparative claims precise: when does a staircase $D(n)$ imply a minimum-description-length bias? Can *flatness* in packing τ (e.g., hyperuniform small-k suppression) be mapped to flat minima or sparsity in representation learning in a provable way?

9.3 Designing computation by choosing distributions, not rules

Design recipe (actionable):

1. Pick τ & $\mathcal{O}$. Decide what "computational regularity" you want (e.g., mobile packets, absorbing cycle, hyperuniform spectrum). Declare the observables and metric D.
2. Select a distribution family $\mathcal{D}(\theta)$.
    - Sizes: (ρ, ϕ_ℓ) for bidisperse; polydispersity σ_r.
    - Shapes: dopant type and fraction ϕ_{shape}, orientation variance.
3. Sweep & score. Run quasi-static PBC protocols; compute the automaton index $\mathcal{A}$ (rule sharpness, packet mobility, TE, entropy rate) alongside endpoint D.
4. Pick operating point. Choose $\theta^\star$ on the automaton ridge (high $\mathcal{A}$, acceptable endpoint D).
5. Validate logic. Extract motif tables $p(y = 1 \mid m)$; demonstrate glider collisions (annihilation/reflect/bind) and reproducible staircases.
6. Robustness. Stress-test against step size, solver tolerance, N, and small stochasticity; retain behavior if $\mathcal{A}$ and Dstatistics stay within bands.

Programming patterns:

- Signal lines: tune $\rho \in [1.2,1.35]$, $\phi_\ell \in [0.15,0.30]$to create mobile defect "wires."
- Logic junctions: add sparse shape dopants ($\phi_\triangle \sim 0.15 - 0.25$) to template Y/T/χ motifs; measure branching/reflection probabilities—your primitive gates.
- Memory/absorption: switch to weak cyclic drive below threshold to settle into absorbing τ (no events)—a reset or "halt" state.
- Spectral cleaning: aim for hyperuniform τ when you want globally smooth outputs; verify small-k suppression without ever optimizing $S(k)$ directly.

From lab to theory and back.

- Experimental realizations: quasi-2D granular beds, vibrated colloids, optical-tweezer disks/triangles, microfluidic emulsions. Distributions are easy to control (size polydispersity, dopant fraction), and PBC can be approximated with rings/annuli.
- Software artifacts to share: reference implementations for (i) event-driven quasi-static protocols with PBC, (ii) flip rasters & avalanche segmentation, (iii) defect trackers, (iv) TE estimators, (v) modular D library (KL on | ψ_6 |/motifs/$S(k)$).

- Ablations: (a) remove dopants (does logic vanish?), (b) randomize motifs while preserving degree (does rule sharpness drop?), (c) swap τ (does admissibility still hold?).

Why this matters.
By relocating *programmability* from hand-coded rules to distributions under a universal admissibility principle, we get: (i) a falsifiable theory (measure ΔD and the staircase), (ii) a design lever that's manufacturable, and (iii) a bridge to learning-by-compression that is conceptually tight yet honest about mechanism. The outlook is a general engineering discipline of materials that compute by alignment—choosing $\mathcal{D}$ to sculpt the induced logic, reading out the computation in flips, spectra, and defect traffic, and certifying correctness by monotone descent of $D(q \parallel \tau)$.

References

1. Liu, A. J., & Nagel, S. R. (1998). Jamming is not just cool any more. *Nature*, 396, 21–22.
2. O'Hern, C. S., Langer, S. A., Liu, A. J., & Nagel, S. R. (2003). Jamming at zero temperature and zero applied stress: The epitome of disorder. *Phys. Rev. E*, 68, 011306.
3. Torquato, S., & Stillinger, F. H. (2003). Local density fluctuations, hyperuniform systems, and order metrics. *Phys. Rev. E*, 68, 041113.
4. Torquato, S. (2018). Hyperuniform states of matter. *Rev. Mod. Phys.*, 89, 031002.
5. Corté, L., Chaikin, P. M., Gollub, J. P., & Pine, D. J. (2008). Random organization in periodically driven systems. *Nature Physics*, 4, 420–424.
6. Keim, N. C., Paulsen, J. D., Zeravcic, Z., Sastry, S., & Nagel, S. R. (2019). Memory formation in matter. *Rev. Mod. Phys.*, 91, 035002.
7. Falk, M. L., & Langer, J. S. (1998). Dynamics of viscoplastic deformation in amorphous solids. *Phys. Rev. E*, 57, 7192–7205.
8. Maloney, C. E., & Lemaître, A. (2006). Amorphous systems in athermal, quasistatic shear. *Phys. Rev. E*, 74, 016118.
9. Argon, A. S. (1979). Plastic deformation in metallic glasses. *Acta Metallurgica*, 27, 47–58.
10. Kosterlitz, J. M., & Thouless, D. J. (1973). Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C*, 6, 1181–1203.
11. Halperin, B. I., & Nelson, D. R. (1978). Theory of two-dimensional melting. *Phys. Rev. Lett.*, 41, 121–124.
12. Young, A. P. (1979). Melting and the vector Coulomb gas in two dimensions. *Phys. Rev. B*, 19, 1855–1866.
13. Weaire, D., & Hutzler, S. (1999). *The Physics of Foams*. Oxford University Press.
14. Aurenhammer, F. (1991). Voronoi diagrams—A survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23, 345–405.
15. Chaikin, P. M., & Lubensky, T. C. (1995). *Principles of Condensed Matter Physics*. Cambridge University Press.
16. Bi, D., Lopez, J. H., Schwarz, J. M., & Manning, M. L. (2015). A density-independent rigidity transition in biological tissues. *Nature Physics*, 11, 1074–1079.

17. Olsson, P., & Teitel, S. (2007). Critical scaling of shearing rheology at the jamming transition. *Phys. Rev. Lett.*, 99, 178001.

18. Wyart, M., Silbert, L. E., Nagel, S. R., & Witten, T. A. (2005). Effects of compression on the vibrational modes of marginally jammed solids. *Phys. Rev. E*, 72, 051306.

19. Makse, H. A., Johnson, D. L., & Schwartz, L. M. (2000). Packing of compressible granular materials. *Phys. Rev. Lett.*, 84, 4160–4163.

20. Durian, D. J. (1995). Foam mechanics at the bubble scale. *Phys. Rev. Lett.*, 75, 4780–4783.

21. Hexner, D., Urbani, P., & Zamponi, F. (2018). Can a large packing be hyperuniform? *Phys. Rev. Lett.*, 120, 068001.

22. Berthier, L., & Biroli, G. (2011). Theoretical perspective on the glass transition and amorphous materials. *Rev. Mod. Phys.*, 83, 587–645.

23. Dahmen, K. A., Ben-Zion, Y., & Uhl, J. T. (2011). A simple analytic theory for the statistics of avalanches in sheared granular materials. *Nature Physics*, 7, 554–557.

24. Denny, R. A., Reichhardt, C., & Reichhardt, C. J. O. (2003). Dynamical phases of vortices in superconductors with periodic pinning. *Phys. Rev. E*, 67, 056110. (Analogy for driven, absorbing-like states)

25. Thurston, W. P. (2005). *The Geometry and Topology of Three-Manifolds* (Chapter on Circle Packing). Princeton lecture notes.

26. Stephenson, K. (2005). *Introduction to Circle Packing: The Theory of Discrete Analytic Functions*. Cambridge University Press.

27. Hexner, D., & Levine, D. (2015). Hyperuniformity of critical absorbing states. *Phys. Rev. Lett.*, 114, 110602.

28. Tjhung, E., & Berthier, L. (2015). Criticality and universality in the reversible-irreversible transition in sheared suspensions. *J. Stat. Mech.*, 2015, P03003.

29. Paulsen, J. D., Keim, N. C., & Nagel, S. R. (2014). Multiple transient memories in sheared suspensions. *Phys. Rev. Lett.*, 113, 068301.

30. Mouyen, M., et al. (choose a recent granular hyperuniformity/structure-factor study relevant to your experiments, if applicable).